R E P O R T . R E S U M E S

ED 013 782 SP 001 256

THE EFFECTS OF SELF-FEEDBACK AND REINFORCEMENT ON THE ACQUISITION OF A TEACHING SKILL.

BY- MCDONALD, FREDERICK J. AND OTHERS

PUB DATE 66

EDRS PRICE MF-$0.25 HC-$1.28 32P.

DESCRIPTORS- *DISCRIMINATION LEARNING, *FEEDBACK, *REINFORCEMENT, *STATISTICAL ANALYSIS, STUDENT PARTICIPATION, *STUDENT TEACHER RELATIONSHIP, TABLES (DATA), TEACHER INFLUENCE, *TEACHER INTERNS, VIDEO TAPE RECORDINGS

TO TEST THE RELATIVE EFFECTIVENESS OF 3 TRAINING PROCEDURES FOR ACQUIRING A TEACHING SKILL, EACH APPLYING REINFORCEMENT PRINCIPLES, STANFORD TEACHER INTERNS WERE VIDEOTAPED ON 4 OCCASIONS DURING THE FIRST 20 MINUTES OF CLASS. EACH INTERN SAW A VIDEOTAPE PLAYBACK WITHIN 3 DAYS (NEW LESSONS WERE VIDEOTAPED WITHIN 2 DAYS AFTER PLAYBACK). REINFORCEMENT TRAINING WAS THE VARIABLE, WITH THE PREDICTED ORDER OF EFFECTIVENESS GOING FROM SELF-ADMINISTERED FEEDBACK TO EXPERIMENTER-ADMINISTERED FEEDBACK, TO EXPERIMENTER ADMINISTERED FEEDBACK WITH CUE DISCRIMINATION TRAINING. A CONTROL GROUP AND THE FOLLOWING 3 EXPERIMENTAL GROUPS FORMED WERE--(1) SELF-FEEDBACK GROUP (S-F) INSTRUCTED IN THE EDUCATIONAL RELEVANCE OF INCREASING STUDENT PARTICIPATION, DEFINED IN TERMS OF PUPIL-PARTICIPATION RESPONSES (PPR), WITH EMPHASIS ON IMMEDIATE REWARD OF PPR'S (PLAYBACKS VIEWED ALONE, EXAMPLES AND A RATING CHART PROVIDED), (2) REINFORCEMENT-ONLY GROUP (R) RECEIVED THE SAME INSTRUCTIONS AS S-F (VIEWED PLAYBACKS WITH AN EXPERIMENTER, WHO REINFORCED INTERNS' REINFORCEMENT OF PPR'S), (3) REINFORCEMENT AND DISCRIMINATION TRAINING GROUP (R AND D) RECEIVED THE SAME INSTRUCTIONS AS S-F (EXPERIMENTER SERVED THE SAME FUNCTION AS FOR R AND ALSO GAVE DISCRIMINATION TRAINING INCLUDING CUES, SUGGESTIONS, AND POSSIBLE EFFECTS). RESULTS WERE ANALYZED BY ANALYSIS OF VARIANCE, T TESTS AND MULTIPLE REGRESSION ANALYSIS. PREDICTIONS WERE BORNE OUT. SUGGESTIONS FOR FUTURE STUDIES AND FOR IMPROVEMENT OF SELF-FEEDBACK ARE INCLUDED. (AF)

~~Experiment I~~

THE EFFECTS OF SELF-FEEDBACK AND REINFORCEMENT ON THE ACQUISITION OF A TEACHING SKILL

Frederick J. McDonald, Dwight W. Allen, and Michael E. Orme

(Stanford University)

The experiment described here applies well known principles of reinforcement theory of a training problem. The training paradigm involves applying a reinforcer to an emitted response. The prediction is that the rate of responding will increase.

Two aspects of the training procedures do, however, depart from those typically used in reinforcement studies. First, the reinforcer is not given while the learner is actually emitting the response. In this study, the subjects were videotaped while emitting complex responses (teaching) some of which were to be reinforced. After the actual behavior sample was collected the subject viewed his performance in the presence of an experimenter. When the desired responses appeared on the videotape, the experimenter reinforced their occurrence. If results similar to those obtained in other studies occur in this situation, reinforcement concepts are widely applicable. Also training procedures or complex skills can be developed which use these principles and concepts in ways directly analagous to the procedures used in laboratory studies where the utility of these concepts has been amply demonstrated.

The second characteristic of the training procedure which departs from the usual laboratory methods is that the behaviors to be learned occur in the context of many other behaviors and are relatively more complex than operants conditioned in laboratory studies. They are more analogous to the kinds of behaviors that have been verbally conditioned on psychotherapy sessions. Here, as in these other complex verbal interactions, it is literally impossible to reinforce every instance of the operant being conditioned. The immediacy of the reinforcement, for similar reasons, is also somewhat variable. Again, if comparable results are obtained, the generalizability of reinforcement concepts is supported.

Since the subjects in this experiment are humans it was also possible to test the efficacy of a cue discrimination procedure. This procedure consisted in pointing out to the subjects those cues to which the reinforced operant was attached. In this way the cue-response chain is clearly indicated which should be facilitating.

Since human subjects are being used, it is also possible that they can reinforce themselves or, more generally, provide their own feedback. This can easily be done with videotape recordings of a behavior sequence since the subject can view himself as a behaving organism. With instructions he can note the presence or absence of the desired response. However, it is dubious if such a procedure is likely to be highly facilitating since the subject may not attend well, may be easily distracted, may be highly subjective in his viewing. The effectiveness of this training procedure remains to be tested, however.

This experiment, then, tests the relative effectiveness of three training procedures, each representing an application of reinforcement principles. The three procedures represent points on a continuum from self-administered feedback or reinforcement to experimenter-administered reinforcement with cue discrimination training.

The prediction is that the order of training effectiveness will be in the same direction, with the self-administered feedback the least effective and the experimenter-administered feedback with cue discrimination training the most effective.

An assumption was made that the behavior being learned by the trainees, reinforcing student participatory behavior, will in turn increase the frequency of this behavior. Although the experiment is not designed to test this assumption, it is supported by reinforcement theory, and correlational data relevant to the assumption is presented here. Unlike the experimental situation, the teacher administers reinforcement as the behavior occurs, approximately more closely, therefore, the operant conditioning paradigm.

EXPERIMENT I: DRAFT ONE

METHOD

General Procedure: Intern teachers were videotaped on four separate occasions during the first 20 minutes of regular classroom lessons. In the intervals between each of these taping sessions they received differential feedback as part of their regular supervision. The treatment or supervision sessions were alike for all subjects in that they viewed videotape playbacks of their earlier teaching performance. The mode and amount of feedback given each intern was varied by manipulating the reinforcement and discrimination training provided by an experimenter.

Before the pretest videotapes were recorded, all subjects were told when they would be taped and were asked to present a discussion-type lesson in which teacher-pupil interaction could be observed. This was not a new or unusual experience for them as they had been frequently exposed to the videotaping-playback-supervision process during the previous three months of the Stanford Intern Program.

Pupils in each of the classrooms were informed by the interns beforehand that the portable TV equipment would be present in the room, and that the cameras would be focused on the teacher, not the class.

Treatments: Mode of feedback, type of reinforcement and amount of discrimination training were varied for four experimental groups of interns.

Controls: Group (C): At the beginning of the first playback session these subjects were given written instructions which suggested that as they viewed subsequent playbacks of themselves, they try to determine their effectiveness in relation to: the aims of the lesson; use of examples; effectiveness of teacher questions; amount of pupil participation; pacing of the lesson, and teacher-pupil rapport. Following this, they viewed the first and all subsequent playbacks alone. E started the machine and left the room. He returned as the tape finished, stopped the machine and told the intern when to expect the next taping and the date for the following playback session. As for all subjects, Group C (controls) viewed playbacks of the preceding lesson within three days of its taping. Lessons were videotaped within two days of each playback session.

Self-Feedback: Group S-F): These subjects followed the same basis schedule as the controls, except that they received a different set of written instructions. The instructions discussed the educational relevance of increasing pupil participation in certain types of lessons; defined pupil participatory responses (PPR) as a clearly observable non-verbal or verbal response that was considered desirable; and provided brief examples of such behavior. It was also suggested that the intern immediately reward a PPR when it occurred as this would tend to increase pupil participation. Examples of teacher responses - both verbal and non-verbal were then provided. Finally, a simple rating chart was attached so that the intern could classify his responses to PPR-s as "teacher rewards"; "teacher ignores" "teacher punishes"; or "can't classify." As with the control group, self-feedback subjects viewed each playback alone.

Reinforcement-Only Subjects: Group (R): These interns received the same written instructions as those in the self-feedback condition. However, E viewed each of the three videotaped playbacks with them, and verbally reinforced all observable instances of these Group (R) interns reinforcing PPR's. Whenever the intern was observed to reinforce a PPR, E responded by saying "Good!" "That's it!" etc. Beyond this, he did not comment upon the intern's teaching performance.

Reinforcement Plus Discrimination Training: Group (R + D): These interns were first given written instructions which were identical to those administered to Group (S-F) and Group (R) subjects. In viewing subsequent playbacks with them, E provided differential reinforcement as in the Reinforcement-Only treatment. In addition, he provided discrimination training. This consisted of pointing out salient cues to which reinforcement should be attached, suggestions related to the immediacy, affect loading and types of reinforcement the teacher could use, and finally, the effects of such behavior upon pupil participation. In general terms, then, it might be said that Group (R+D) subjects received "maximum supervision," and Group (S-F) subjects received "self-supervision."

Experimental playback sessions for all groups were thirty minutes in length. When E viewed playbacks with Group (R) and (R+D) subjects, he reduced the sound momentarily when providing discrimination training, and spoke over the tape when providing reinforcement. Interns in all groups were informed that they could have the tape stopped or reversed and played over again at any point during the playback. Playbacks were stopped occasionally by E when he was working with the supervised groups. However, since the sessions were limited to thirty minutes, only one or two brief stoppages in sessions two and three were possible.

Subjects:

All Stanford intern teachers majoring in English, Social Studies and Mathematics were included in the study. Approximately equal numbers of interns from each of the subject-matter areas were assigned to each group. In addition, groups of interns teaching in the same school were distributed throughout the four groups. In this way, systematic bias due to subject-matter major or pupil characteristics based on socio-economic status was avoided. The major characteristics of the sample are summarized in Table I. See Table I, Experiment I).

The Dependent Variable:

The dependent variable was defined as the relative frequency with which the teacher positively reinforces pupils participatory responses during teacher-pupil interaction in the classroom. PPR's were defined in training sessions with the interns as any desirable or relevant pupil comment, answer or question. For purposes of measurement however, desirability and relevancy were not considered.

The basic strategy in defining the dependent variable involved classifying teacher responses into one of four major response categories. These include positive reinforcement, negative reinforcement, interaction and information giving responses. Pupil responses were also classified and considered in relation to teacher responses. Each of the above response categories are defined in the following discussion. A summary of the classification theme appears in Table 2. (See Table 2, Experiment I).

Teacher Positive Reinforcement: A teacher response was defined as positively reinforcing it, met one of two conditions. First, the response had to immediately follow a PPR. Secondly, it had to be classifiable under one of the following response categories. (1) Teacher Positive Verbal Reinforcement (+ VR): Immediately following a PPR, the teacher uses words and phrases such as 'Good,""Fine!"(2) Teacher Positive Non-Verbal Reinforcement (+NVR). The teacher in responding to a PPR, nods, smiles, leans or moves toward the pupil, or writes the pupil's response on the blackboard. (3) Teacher Positively Qualified (+QR) and (4) Post Hoc Reinforcement (PHR). The teacher emphasizes positive aspects of pupil responses by reorienting class attention to earlier contributions by a given pupil (PHR), or by differentially reinforcing the acceptable components of partially adequate response (+QR).

Teacher Negative Reinforcement: A teacher response was defined as negatively reinforcing it immediately followed by a PPR, and was classifiable as the obverse of one of four types of reinforcement outlined above (-VR; -NVR; -QR).

Teacher-Pupil Responses Independent of Reinforcement Classifications: Certain responses that occur frequently in interaction and yet are not classifiable as some form of reinforcement were included in the definition of the dependent variable. These included information-giving by the teacher, teacher initiated interaction (i.e., questions directed to a given pupil or to the class in general) pupil initiated interaction (i.e., volunteered comments or questions), teacher no response and pupil no response.

In general, the dependent variable included evaluative and informational signals which the teacher may use in the classroom. There was a tendancy to emphasize socially rewarding operants since it could be expected that such behaviors would tend to increase pupil participation.

Both in the training and measurement phases of the study, the PPR was presented as an S^D which served to cue the teacher or rater that a desirable (or classifiable) teacher response was about to occur.

Measurement Procedures: Four videotapes for each intern in each group were analyzed by raters trained for this purpose. Throughout the rating phase of the experiment, they worked on the tapes in a random order so that they neither knew the treatment condition nor the number of the teaching trial of the tape being rated. Operators ran the television equipment and selected tapes using a list of random numbers.

In addition to recording the frequencies of each of the behaviors defined as components of the dependent variable, the raters recorded other relevant behaviors and lesson characteristics as well. Frequencies were recorded for the total number of pupils who responded, the number of responses they emitted, and the sex of each responder. The raters also recorded the length of each videotape to the nearest tenth of a minute, and determined how much time was spent in discussion, group work or individual study.

The Unit of Measurement: In analyzing pupil-teacher interaction, one may record discrete responses, or measure in terms of some unit such as the uninterrupted utterance, or the spoken sentence. Both systems of measurement were used in this study. A single pupil response was defined as an uninterrupted utterance. Raters were trained to define an interruption as a comment or question. "Partial" responses that teachers commonly emit during pupil speech ("um-hum"; "Yes,"; etc.) were not defined as interruptions. All forms of verbal and non-verbal reinforcement were scored in terms of discrete responses. For example, if the teacher said, "Good!"; "Good"; "That's fine!", the rater coded all three operants.

Training of Raters: Eight raters were initially given intensive training on intern videotapes. Once they had achieved at least 90% interrater agreement on all of the major response categories, and better than 95% agreement on teacher reinforcement responses, the analysis of experimental tapes was begun. Reliability was maintained throughout the analysis by scheduling frequent joint rating sessions where raters checked the percentage of agreement and referred to definitions of relevant responses so that systematic rating biases would not develop. Neither ratings taken during training, nor those produced in the joint sessions were used in the statistical analysis of results.

The ratings upon which the reliability coefficients reported in Table 3 are based, were acquired in the following way. As each block of 30 or 40 tapes were completed, (a total of 269 were actually rated), each of the six raters who did the bulk of the rating then rated a given tape. This was done without the rater's knowledge. Eight tapes, two from each trial, and two from each group were rated by all six raters in this way. As can be seen, interrater agreement is high. (See Table 3, Experiment I)

A certain amount of data was lost between initial videotaping and the final statistical analysis. Some tapes were technically poor, and could not be rated. Some tapes were inevitably less than the required 20 minutes, and were also omitted. However, if a given tape was over 15 but less than 20 minutes, the obtained ratings were prorated to bring them up to the 20 minutes criterion. Of the initial 284 videotaped lessons then, 25 were omitted at the outset, and 51 tapes were prorated before statistical operations were performed. The T statistic was used to determine whether or not certain cells in the matrix were biased by the inclusion of a disproportionate number of adjusted tapes. T was non-significant. Omitted and prorated tapes are shown by group and trial in Table 4. As can be seen, they are scattered throughout groups and trials. (See Table 4, Experiment I)

RESULTS

Three types of analysis were performed upon the data. Analyses of covariance were employed to test for the significance of differences between each of the groups (treatment differences). In addition, T tests were used to determine the significance of differences within a given group from one trial to the next (training differences). Finally, a multiple regression analysis was performed on all of the major response categories in order to determine significant relationships among these teacher-pupil behaviors.

Treatment Differences: Positive teacher reinforcement constitutes the major response category of the dependent variable. Using trial one scores as covariants, the groups were found to be significantly different from each other. The data in Table 5 summarize these results, and show that the differences were significant for both positive verbal ($p=.001$, .005 and .025 for trials 2, 3 and 4) and non-verbal reinforcement ($p.= .025$ and .005 for trials 2 and 4). (See Table 5, Experiment I).

When the two types of positive reinforcement are taken together, it can be seen that the R+D group outperformed all other groups. These relationships are illustrated in Figure I. (See Figure 1, Experiment I). Positive verbal and non-verbal teacher reinforcements are presented in terms of unadjusted mean frequencies for all trials and groups.

Negative verbal and non-verbal teacher reinforcement occurred infrequently throughout the groups. Table 6 shows that all three of the experimental groups consistently emitted fewer negatively reinforcing responses over trials while control group responses increased. These differences appear to be fairly stable, but do not reach an acceptable level of significance. (See Table 6, Experiment I) Figure 2 illustrates group trends in negative reinforcement. The data here are also based on unadjusted treatment means. (See Figure 2, Experiment I)

Training Differences: Table 7 summarizes the significance levels obtained when within-group treatment means were compared using the T statistic (Winer, 1962, pp. 65-60). As can be seen, Group (R+D) interns significantly increased their rate of positive reinforcement by trial 2 ($p<.01$). (See Table 7, Experiment I) They increased from a base rate of reinforcing

approximately 60% of all PPR's to a rate of 76% by trial 2. In trial 4 the rate dropped to 67%, and this combined with considerable variation within the group, produced before and after treatment differences that were non-significant (p=.10). However, when trial 1 versus trial 4 differences were tested on the assumption that the population variances were unequal, significance well beyond the .05 level was obtained (required .05 = 2.13; obtained 7 = 4.19, Ferbuson, 1959, pp. 143-145). Group (R+D) subjects also increased their rate of positive non-verbal reinforcement from trial 1 to 2 and following two treatment sessions, ended (p = .10) to use less negative verbal reinforcement.

Increases in mean positive verbal reinforcement and a concomitant drop in negative reinforcement can be most clearly seen in Group (R) subjects. The higher significance revels result from considerably less within-group variation.

The control group showed no significant within-group shifts in reinforcement from trial 1 to trial 4. Note that the self-feedback subjects tended to increase their negative verbal reinforcement rate throughout treatment.

Pupil Responses: What are the probable effects of the above types of teacher training on pupil behavior? To answer this question, total pupil responses and relevant component responses were analyzed. A summary of the analyses of covariance (Table 8) performed on Total Pupil responses with trial one scores as covariates, shows that shifts in pupil responses closely allowed concomitant shifts in teacher positive reinforcement. (See Table 8, Experiment I) while the increase inrthe Group (R) pupil sample appears short-lived, Group (R+D) pupils maintain significantly higher response levels in trial 4 (p.005) as well as for trials 2 (p=.001) and 3 (p=.01). These data are illustrated in Figure 3. (See Figure 3, Experiment I) Note that while the control group initially showed a higher mean frequency of responses it dropped slightly over four trials. Group (R) and (R+D) pupils increased from trial 1 to trial 4.

The increase in total pupil responses immediately leads one to ask whether they are due to increased teacher positive reinforcement, or perhaps more simply, a function of increased questioning by the teacher. While it is clear that differential feedback and reinforcement affected teacher behavior, it does not necessarily follow that increased positive reinforcement as defined here had an affect on pupil behavior. A comparison of volunteered pupil responses as opposed to teacher-solicited pupil responses is relevant. If increased pupil responses were largely due to increased questioning by the teacher then one would detect an increase in directly solicited answers, and a decrease or no change in pupil volunteered statements and questions. This does not appear to be the case. F ratios based covariance analyses of Pupil Volunteered Statements (Table 9, Experiment I) were significant for trials 2 (p =.01) and 4 (p=.01). Group (R+D) subjects after showing sharp gains off somewhat in trial 3, but rose again in trial 4. Note however that control group in frequencies followed a similar pattern while those for groups (R) and (S-F) moved in opposite direction. Figure 4 illustrates these trends. (See Figure 4, Experiment I).

Teacher-solicited pupil responses, while obviously related to total pupil responses not entirely account for the increases noted. That this is so, can clearly be seen in figures 5 and 6. (See Figures 5 and 6, Experiment I)

Figure 5 illustrates the relationships between teacher specified responses and pupil responses in relation to all pupil responses. It is a graphical representation of the relevant relation coefficients reported in the next section. Mean treatment frequencies of those responses directly solicited by the teacher (DSI and DSG-1) and pupil volunteered statements and questions (V and V?) were summed for each trial. This sum was then divided by total pupil volunteered responses (V+V?) so that the proportion of pupil volunteered responses to teacher solicited responses could be determined for each trial.[1] Note that from trials 1 to 4, an increasing percentage of volunteered responses contributes to the combined totals in Groups (R) and (R+D) while the control group shows smaller increases. Self-feedback subjects show a distinctly different pattern. This decrement is consistent with their performance in terms of teacher reinforcement.

---

1 $\frac{(V + V?)}{(DSI + DSG\text{-}1) + (V + V?)}$

In Figure 6, teacher specified (DSI + DSG-I) and pupil volunteered responses (V+V?) are presented for the controls and Group (R+D). The frequencies plotted are based on unadjusted treatment means.

Finally, it is of interest to consider inter-correlations between various response categories of the dependent variable, and total pupil responses. The intercorrelation matrix for these relationships is presented in Table 10. (See Table 10, Experiment I). These correlational data are consistent with the earlier reported results. Teacher positive reinforcement was found to be significantly related to total pupil responses (0.50) and to volunteered pupil statements (0.46). Teacher specified questions were also significantly related to total pupil responses (0.56). However, when they are considered in relation to volunteered pupil responses, the relationship is very slight (+0.07), and negative (-0.13) when we consider volunteered questions in relation to teacher specified questions which are directed to the class as a whole.

An initially surprising relationship obtains between total pupil responses and negative verbal reinforcement by the teacher (+0.45). One interpretation would be that negative verbal reinforcement produces pupil attention. In addition, it also probably has feedback value -- it will be recalled that "No" and "Wrong" responses by the teacher were included in the negative reinforcement response category.

Discussion of Results

It is clear that feedback conditions proved to be the most effective training arrangement. Perhaps of greatest interest are what appear to be these variations in feedback which are most effective. Clearly, adding cue discrimination to the training method substantially improves the procedure. This procedure is, however, the most "costly" in that it requires the active involvement of the experimenter to describe salient cues and to suggest ways of reinforcing participating behavior that the subject could use.

Variations in the effectiveness of this procedure might occur when a variety of experimenters are used. Informal observations suggest that not all trainees responded equally positively to this condition. Trainee characteristics probably interact with experimenter characteristics, and such interactions probably influence differentially the effects of feedback and cue discrimination training.

Equally interesting is the relative ineffectiveness of the self-feedback condition, attractive because it is the least costly procedure. This method is probably ineffective because the desired response is not adequately cued. Even if trainees had a limited response repertoire of reinforcing responses, they still could have used them consistently if somewhat monotonously. But, the rate increase is not likely to occur if the trainee does not "know" when to emit the desired response. Both of the other feedback conditions cue as to appropriate response in some form.

This self-feedback condition might be improved by introducing some cueing procedures. Or, a combination of viewing models and self might be effective. The results obtained in this experiment indicate only that a limited kind of self-viewing, presumed to be a self-feedback condition and designed to be so, is not highly effective in producing behavior change.

The results of this experiment suggest that the operant conditioning model may be extended to situations in which the learner is not actually behaving but merely watching his performance after the actions have occurred. This extrapolation, if further substantiated, greatly increases the application of this particular paradigm. However, further research must also be directed to an analysis of the viewing conditions -- characteristics of the persons viewing, time interval between enacting and viewing, the kind of behavior being reinforced, and similar conditions which might reasonably be expected to enhance or to emit the effects of the reinforcement procedure.

Also of theoretical interest is the possibility of modifying the operant conditioning paradigm by instructing subjects. This procedure shortens the time and cost of shaping the desired behavior through a series of successive approximations. The learning paradigm of this arrangement needs explication and analysis.

T A B L E 1

(Exp. I)

MEAN AGE AND FREQUENCY DISTRIBUTIONS FOR SEX,

GRADE-LEVEL TAUGHT, AND SUBJECT-MAJOR FOR

EACH EXPERIMENTAL GROUP

N = 71

Experimental Group	Mean Age	Sex		Grade Level		Subject-Major		
		Male	Female	9-10	11-12	English	Soc. St.	Math.
Group 1 Controls (N = 18)	24.9	4	14	13	5	9	4	5
Group 2 Self-Feedback (n = 18)	24.6	5	13	13	5	7	7	4
Group 3 Reinforcement only (N = 18)	24.4	5	13	11	7	5	9	4
Group 4 Reinforcement + Discrimination Training (N = 17)	24.0	3	14	12	5	8	8	1

T A B L E 2

(Exp. I)

THE DEPENDENT VARIABLE DEFINED IN TERMS

OF SELECTED RESPONSE CATEGORIES

Teacher Responses		Teacher - Pupil Interaction	
Positive Reinforcement	Negative Reinforcement	Pupil Initiated	Teacher Initiated
+ Verbal Reinforcement (+VR) + Non-Verbal Reinforcement (+NVR) + Qualified Reinforcement (+QR) Post Hoc Reinforcement (PHR)	- Verbal Reinforcement (-VR) - Non-Verbal Reinforcement (-NVR) -Qualified Reinforcement (-QR)	Volunteer Comments (V) Volunteer Questions (V?) No Response (PNR)	Direct question to one individual (DSI) Direct question to Group, then Individual Specified (DSG-1)
Teacher No Response (TNR)		Teacher question directed to the group, followed by volunteered pupil response (DSG-V)	

T A B L E 3

(Exp. I)

INTERRATER RELIABILITY FOR SIX RATERS

ON THE MAJOR RESPONSE CATEGORIES

OF THE DEPENDENT VARIABLE

Response Category	Kendall Coefficient of Concordance: W
Positive Verbal Reinforcement	1.00
Positive Non-Verbal Reinforcement	0.99
Negative Verbal Reinforcement	1.00
Negative Non-Verbal Reinforcement	0.87
Total Pupil Responses	0.97

T A B L E 4

(Exp. I)

NUMBER OF VIDEOTAPES OMITTED

AND PRORATED BEFORE STATISTICAL ANALYSIS

	Trial 1		Trial 2		Trial 3		Trial 4		TOTALS	
Group	P	O	P	O	P	O	P	O	Prorate	Omit
1 (C)	2	2	2	2	1	0	2	1	7	5
2 (S-F)	3	2	4	1	3	0	1	2	11	5
3 (R)	6	1	1	1	5	1	6	4	18	7
4 (R + D)	4	4	2	2	4	1	5	1	15	8
TOTALS	15	9	9	6	13	2	14	8	51	25

T A B L E 5

(Exp. I)

SUMMARY OF THE ANALYSES OF COVARIANCE

FOR THE EXPERIMENTAL GROUPS, WITH TRIAL ONE SCORES

AS COVARIANTS AND TEACHER POSITIVE VERBAL AND NON-VERBAL

REINFORCEMENTS AS THE DEPENDENT VARIABLES

	Group 1 (C)	Group 2 (S-F)	Group 3 (R)	Group 4 (R+D)	df	F
	Adjusted Means and Standard Errors					
Trial 2 verbal	30.51 4.4678	30.18 4.3160	45.18 4.2841	60.93 4.2532	3/60	11.884***
Trial 2 non-verbal	11.50 3.1517	9.39 3.0566	13.54 2.9530	22.13 2.9556	3/60	3.482*
Trial 3 verbal	22.16 4.6586	25.11 4.3512	40.53 4.3272	53.12 4.3065	3/59	10.846**
Trial 3 non-verbal	13.69 3.2773	7.11 3.0845	12.44 2.9739	14.7709 2.9929	3/59	1.200
Trial 4 verbal	31.42 5.9130	27.84 5.9015	32.26 6.147	51.67 5.3842	3/48	3.721*
Trial 4 non-verbal	6.26 2.5389	7.56 2.5279	9.56 2.7565	19.33 2.2782	3/48	6.315***

* .025 level of significance
** .005 level of significance
*** .001 level of significance

T A B L E 6

(Exp. I)

SUMMARY OF THE ANALYSES OF COVARIANCE FOR THE EXPERIMENTAL GROUPS, WITH TRIAL ONE SCORES AS COVARIANTS AND TEACHER NEGATIVE VERBAL AND NON-VERBAL REINFORCEMENTS AS THE DEPENDENT VARIABLE

	Group 1 (C)	Group 2 (S-F)	Group 3 (R)	Group 4 (R+D)	df	F
	Adjusted Means and Standard Errors					
verbal Trial 2	5.15 1.4149	1.88 1.3759	5.53 1.3392	5.52 1.3287	3/60	1.663[a]
non-verbal	2.64 0.6700	1.61 0.6515	1.54 0.6301	1.09 0.6304	3/60	0.990
verbal Trial 3	4.74 1.5257	5.37 1.4327	3.87 1.3940	6.06 1.3838	3/59	0.446
non-verbal	1.51 0.5384	1.07 0.5054	0.58 0.4897	1.81 0.4899	3/59	1.184
verbal Trial 4	6.24 1.5001	3.33 1.5041	2.59 1.6400	3.25 1.3518	3/48	1.148
non-verbal	1.57 0.5060	0.94 0.5127	0.84 0.5491	0.75 0.4623	3/48	0.535

[a] .250 level of significance

T A B L E 7

(Exp. I)

SIGNIFICANCE OF DIFFERENCES FOR THE FOUR EXPERIMENTAL GROUPS FROM TRIAL TO TRIAL, ON MAJOR RESPONSE CATEGORIES OF THE DEPENDENT VARIABLE*

Response Category	Group 1 (Control)	Group 2 (S-F)	Group 3 (R)	Group 4 (R+D)
	Direction of Difference from Trial to Trial and Level of Significance			
+VR	NS	NS	$T_1 < T_2$ (.01) $T_4 < T_2$ (.05)	$T_1 < T_2$ (.01) $T_1 < T_4$ (.10)
+NVR	NS	NS	NS	$T_1 < T_2$ (.10)
-VR	NS	NS	NS	NS
-NVR	NS	$T_1 > T_3$ (.10) $T_1 > T_4$ (.10)	$T_4 < T_1$ (.05) $T_3 < T_1$ (.05)	$T_4 < T_1$ (.10)

*Significance levels were tested by the T statistic for comparisons among treatment means (Winer, 1962, pp. 65-70).

T A B L E 8

(Exp. I)

SUMMARY OF THE ANALYSES OF COVARIANCE FOR THE EXPERIMENTAL GROUPS, WITH TRIAL ONE SCORES AS COVARIANTS AND TOTAL STUDENT RESPONSES AS THE DEPENDENT VARIABLE

	Group 1 (C)	Group 2 (S-F)	Group 3 (R)	Group 4 (R+D)	df	F
	Adjusted Means and Standard Errors					
Trial 2	72.59 5.5354	47.39 5.2296	65.93 5.3014	83.29 5.1161	3/60	8.267***
Trial 3	58.97 6.6169	54.26 6.0530	60.06 6.3019	84.31 5.9179	3/59	5.082**
Trial 4	61.40 6.9336	51.37 7.0353	52.65 7.8020	80.64 6.5346	3/48	3.913*

* .05 level of significance
** .01 level of significance
*** .001 level of significance

T A B L E 9

(Exp. I)

SUMMARY OF THE ANALYSES OF COVARIANCE FOR THE EXPERIMENTAL GROUPS WITH TRIAL ONE SCORES AS COVARIANTS AND TOTAL STUDENT VOLUNTARY STATEMENTS AS THE DEPENDENT VARIABLE

	Group 1 (C)	Group 2 (S-F)	Group 3 (R)	Group 4 (R+D)	df	F
	Adjusted Means and Standard Errors					
Trial 2	13.23 2.4159	5.98 2.3169	12.51 2.2934	17.31 2.2522	3/60	4.184**
Trial 3	10.14 2.3562	8.20 2.1965	10.95 2.1850	15.10 2.1588	3/59	1.786[a]
Trial 4	12.27 2.3098	7.39 2.3056	8.80 2.5336	17.41 2.0725	3/48	4.159**

[a] 250 level of significance
** .01 level of significance

T A B L E 1 0

(Exp. I)

INTERCORRELATION MATRIX FOR THE MAJOR RESPONSE CATEGORIES

OF THE DEPENDENT VARIABLE

Response Category		TPR	DSI	DSG-I	DSG-V	V	+VR	+NVR	-VR	-NVR
		1	2	3	4	5	6	7	8	9
Total Pupil Responses (TPR)	1	1.00	0.56	0.09	0.41	0.56	0.50	0.16	0.45	0.07
Teacher Question to Individual Pupils (DSI)	2		1.00	0.08	0.13	0.07	0.33	0.03	0.22	0.02
Question to Class, then to one Individual (DSG-I)	3			1.00	0.02	-0.13	0.09	0.02	0.02	-0.03
Question to Class, then a pupil Volunteer (DSG-V)	4				1.00	0.14	0.24	0.17	0.11	0.10
Volunteer Statement by Pupils (V)	5					1.00	0.46	0.23	0.20	0.007
Teacher Positive Verbal Reinforcement (+VR)	6						1.00	0.44	0.14	0.007
Teacher Positive Non-Verbal Reinforcement (+NVR)	7							1.00	0.02	0.14
Teacher Negative Reinforcement (-VR)	8								1.00	0.15
Teacher Negative Non-Verbal Reinforcement (-NVR)	9									1.00

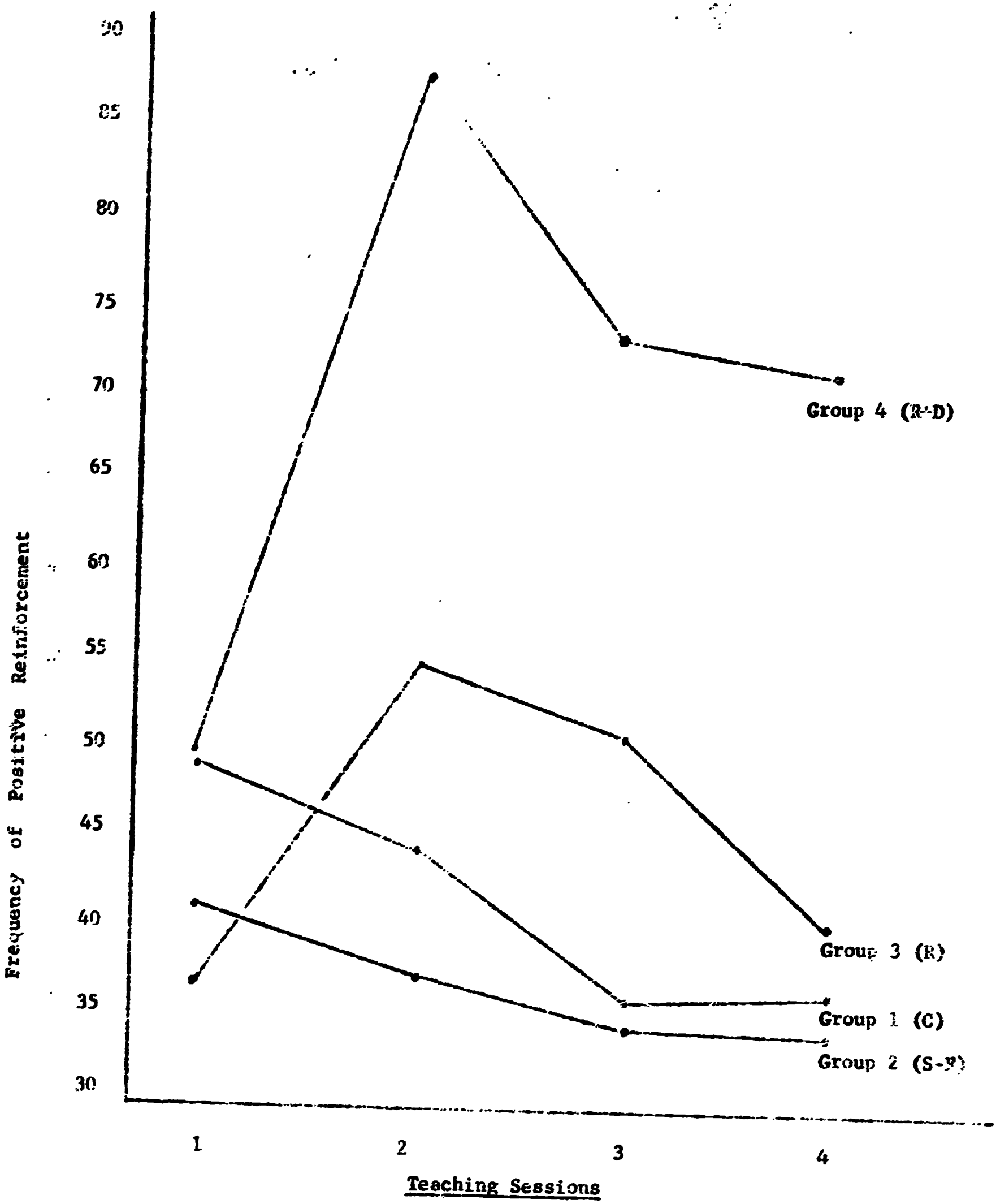

FIGURE L. Unadjusted mean frequencies of teacher positive verbal and non-verbal reinforcement in each of the four treatment groups over four trials.

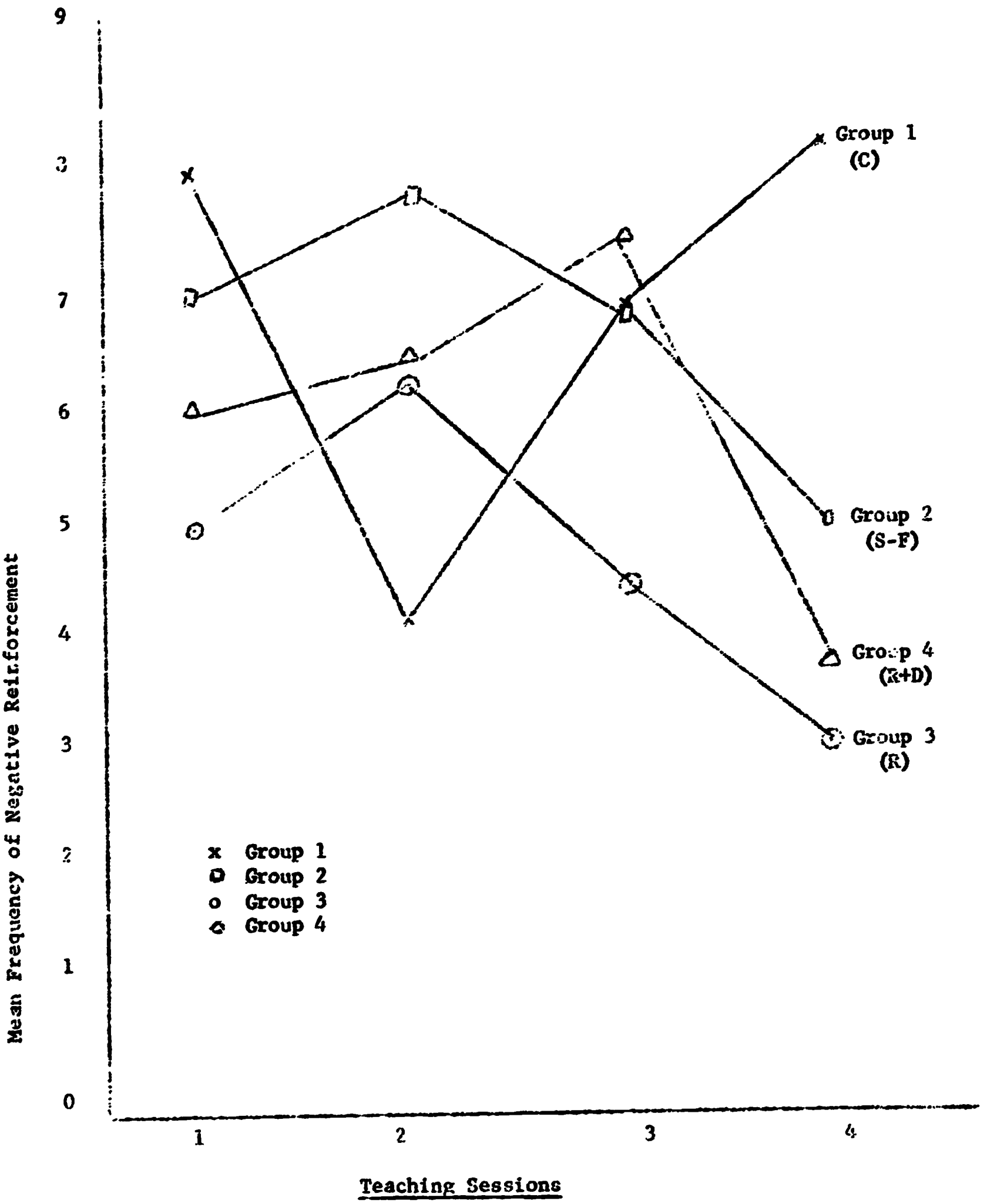

FIGURE 2. Unadjusted mean frequencies of teacher negative verbal and non-verbal reinforcement in each of the four treatment groups over four trials.

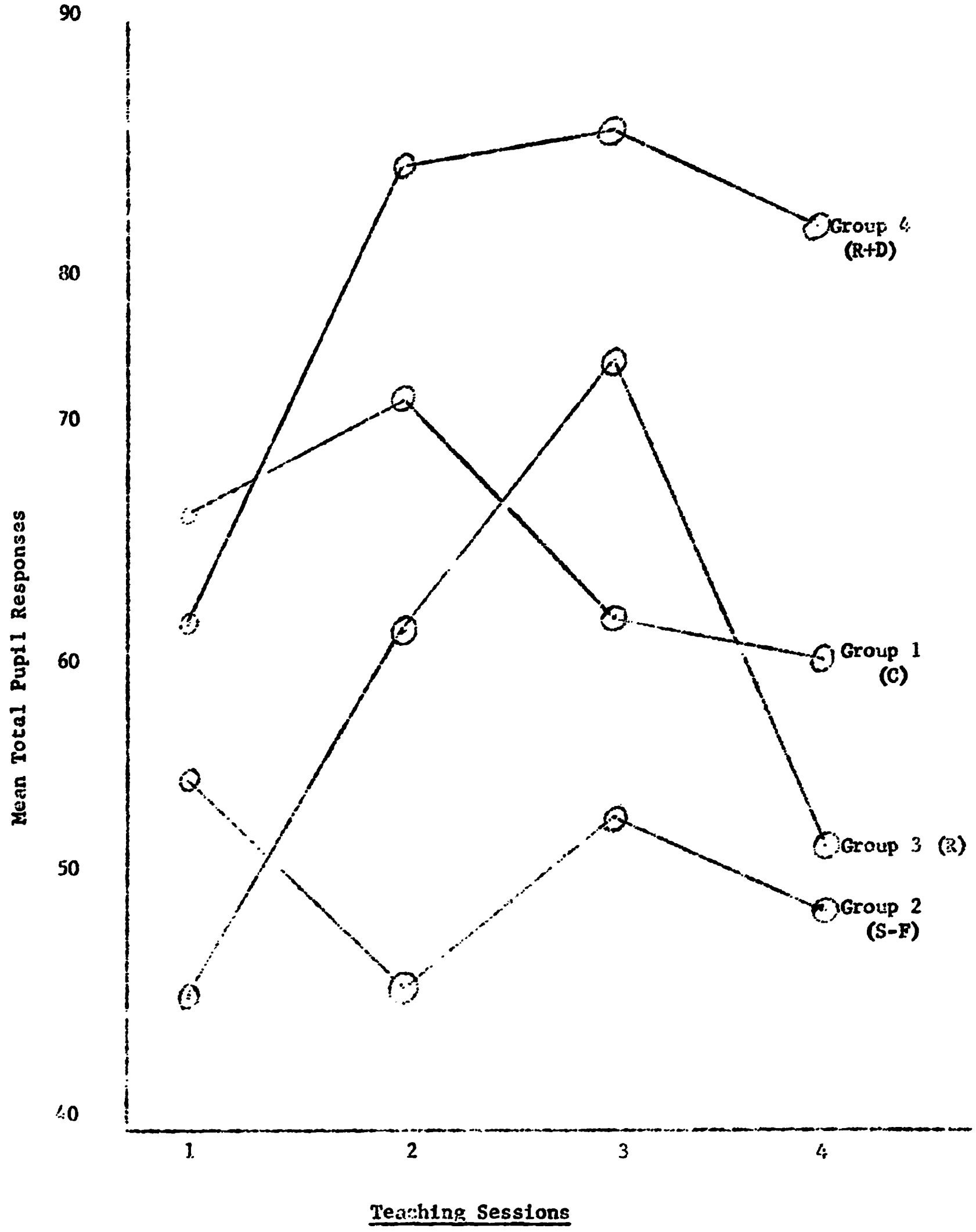

FIGURE 3. Unadjusted mean frequencies of total pupil responses in each of the four treatment groups over four trials.

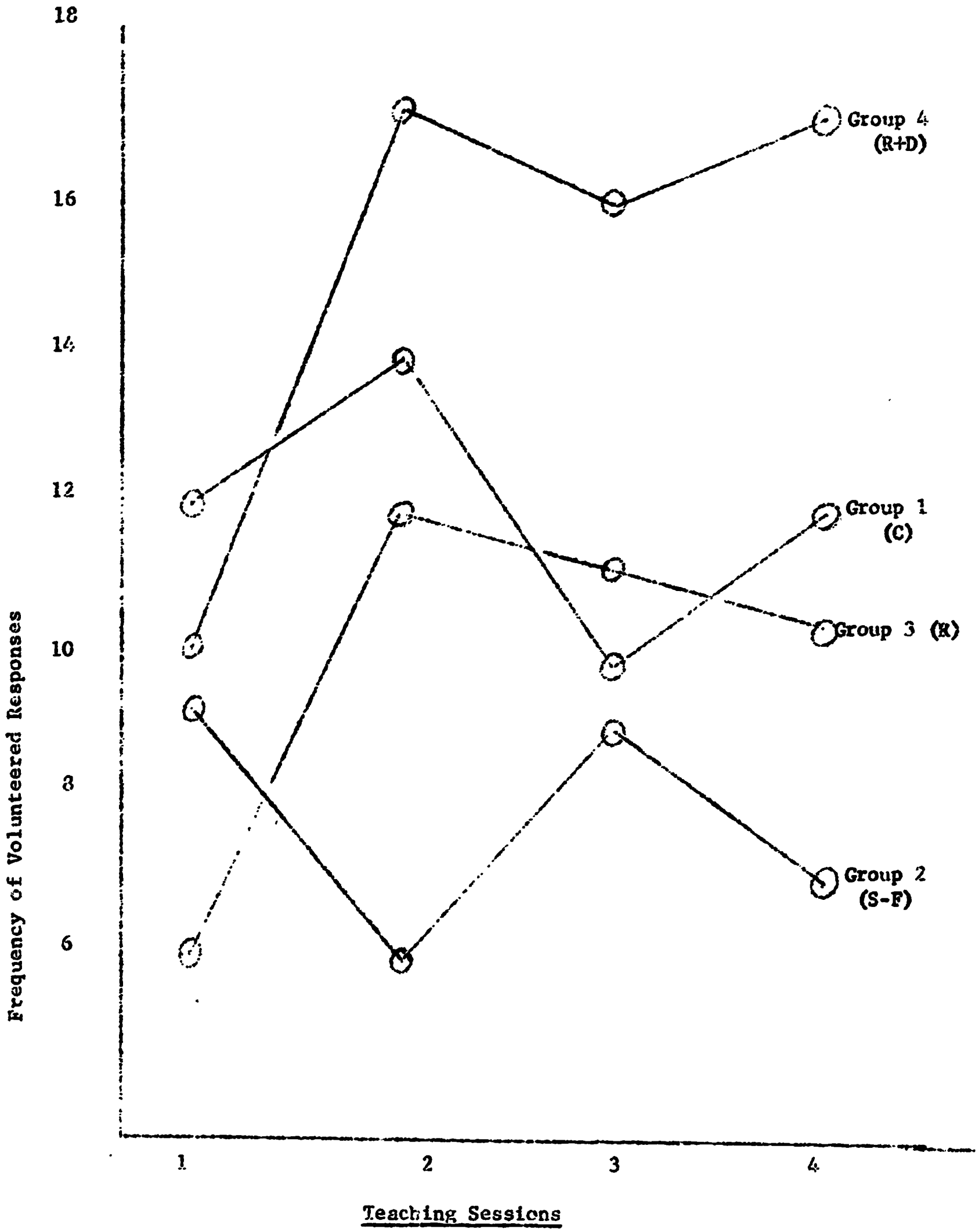

FIGURE 4. Unadjusted mean frequencies of volunteered pupil statements for each of the four treatment groups over four trials.

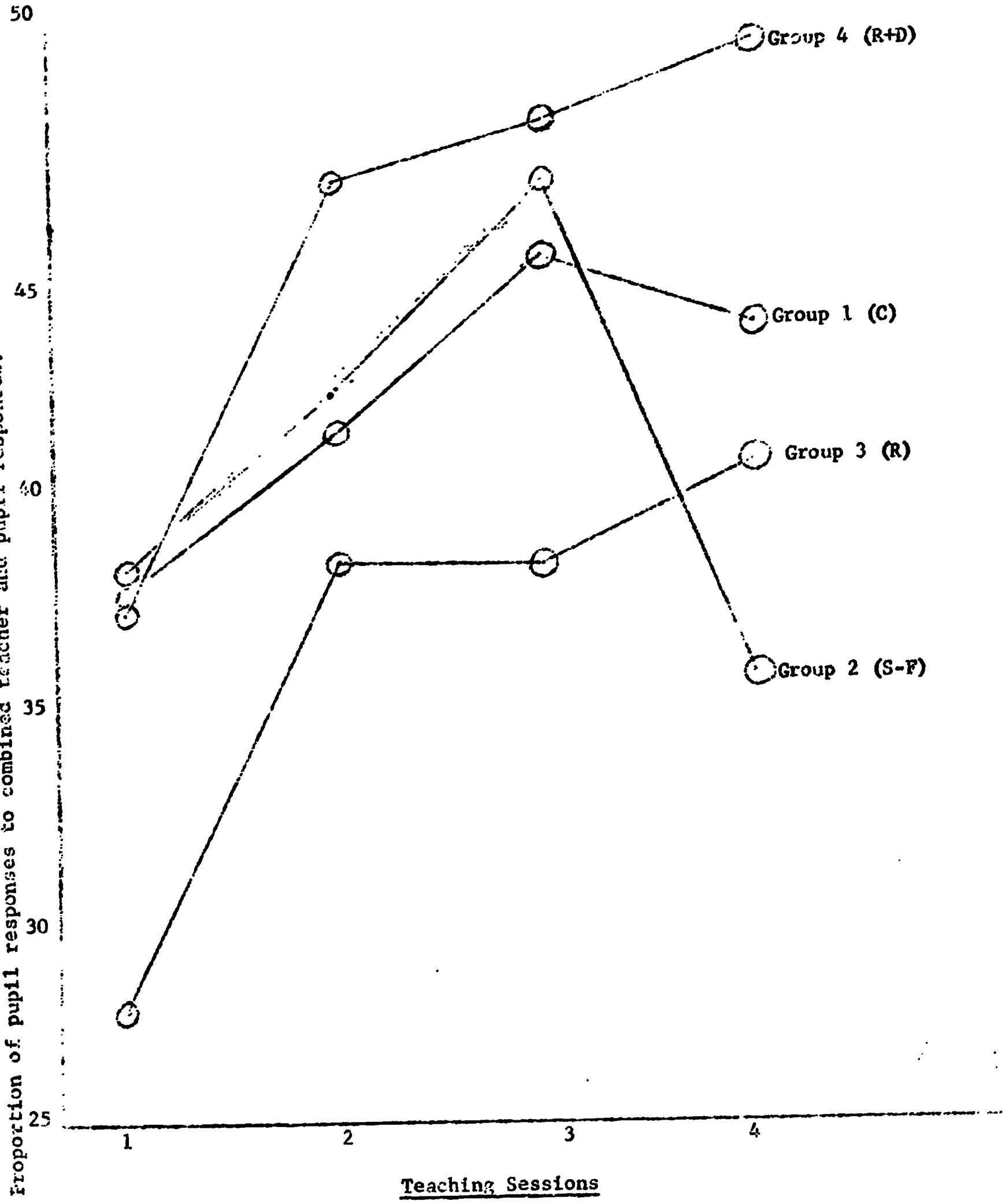

FIGURE 5. Proportion of pupil initiated responses in relation to combined teacher and pupil initiated responses for each of the four treatment groups over four trials.

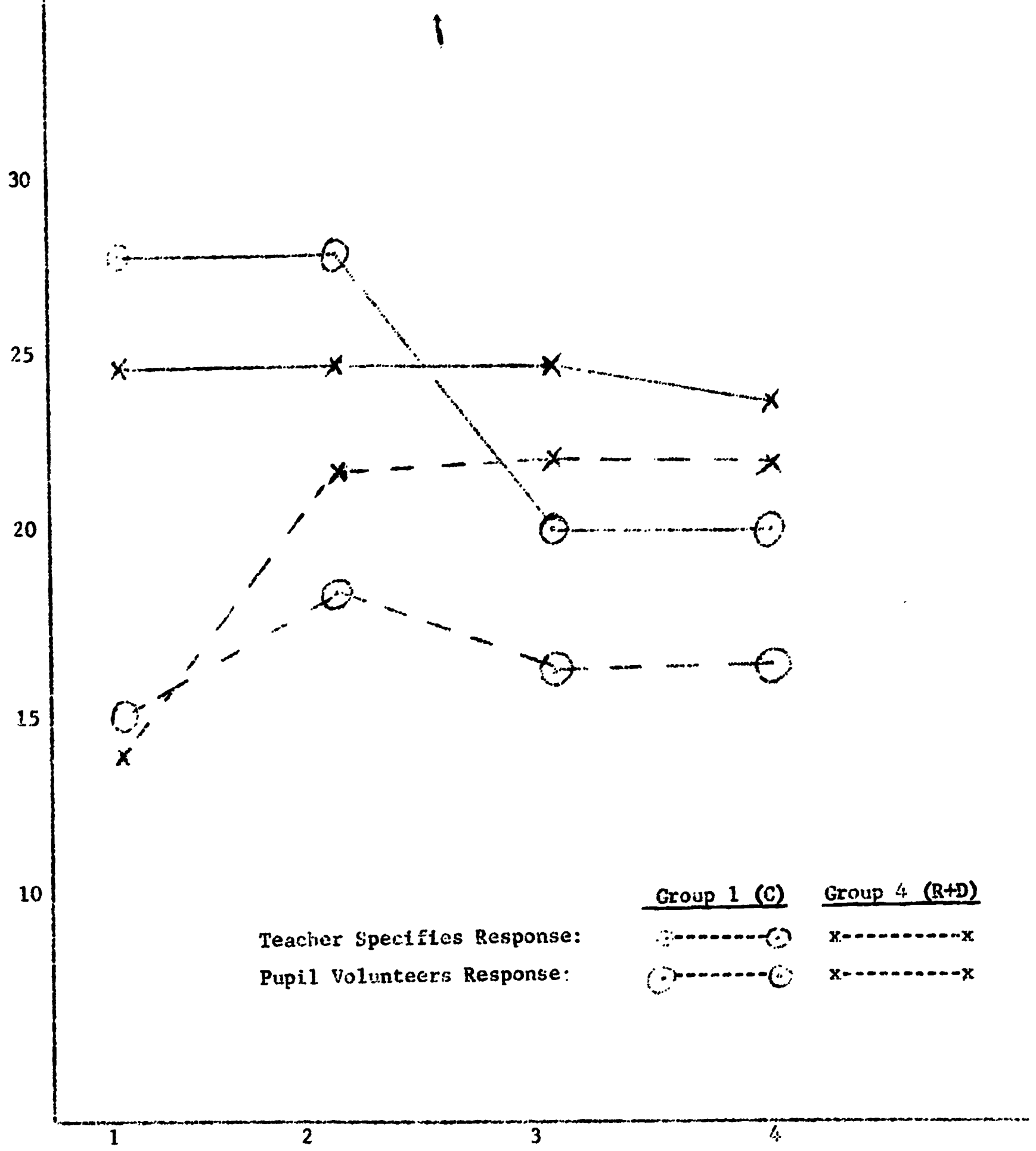

FIGURE 6. Unadjusted mean frequencies of teacher-specified and pupil volunteered responses as they vary from trial one to trial four for the control group and the maximum treatment group.